

A SIMPLE NUMERICALLY STABLE PRIMAL-DUAL ALGORITHM FOR COMPUTING NASH-EQUILIBRIA IN SEQUENTIAL GAMES WITH INCOMPLETE INFORMATION

Elvis Dohmatob

Parietal Team, INRIA / CEA, Université de Paris-Saclay

ABSTRACT

We present a simple primal-dual algorithm for computing approximate Nash-equilibria in two-person zero-sum sequential games with incomplete information and perfect recall (like Texas Hold'em Poker). Our algorithm is numerically stable, performs only basic iterations (i.e. matvec multiplications, clipping, etc.), and no calls to external first-order oracles, no matrix inversions, etc.), and is applicable to a broad class of two-person zero-sum games including simultaneous games and sequential games with incomplete information and perfect recall. The applicability to the latter kind of games is thanks to the sequence-form representation which allows us to encode any such game as a matrix game with convex polytopal strategy profiles. We prove that the number of iterations needed to produce a Nash-equilibrium with a given precision is inversely proportional to the precision. As proof-of-concept, we present experimental results on matrix games on simplexes and Kuhn Poker.

Index Terms— Nash-equilibrium, sequential games, incomplete information, perfect recall, convex optimization

1 Introduction

A game-theoretic approach to playing games strategically optimally consists of computing Nash-equilibria (in fact, approximations thereof) offline, and playing one's part (an optimal *behavioral* strategy) of the equilibrium online. This technique is the driving-force behind solution concepts like CFR [1, 2, 3], CFR⁺ [4] and other variants, which have recently had profound success in Poker. However, solving games for equilibria remains a mathematical and computational challenge, especially in sequential games with imperfect information. In this paper, we propose (our detailed contributions are sketched in subsection 1.3 below and elaborated in section 3) a simple primal-dual algorithm for solving for such equilibria approximately (in a sense to be made precise in Definition 3 below).

1.1 Statement of the problem

The sequence-form representation for two-person zero-sum games with incomplete information was introduced in [5], and the theory was further developed in [6, 7, 8] where it was established that for such games, there exist sparse matrices $A \in \mathbb{R}^{n_1 \times n_2}$, $E_1 \in \mathbb{R}^{l_1 \times n_1}$, $E_2 \in \mathbb{R}^{l_2 \times n_2}$, and vectors $e_1 \in \mathbb{R}^{l_1}$, $e_2 \in \mathbb{R}^{l_2}$ such that n_1 , n_2 , l_1 , and l_2 are all linear in the size of the game tree (number of states in the game) and such that Nash-equilibria correspond to pairs (x, y) of *realization plans* which solve the primal LCP (Linear Convex Program)

$$\begin{aligned} & \underset{(y, p) \in \mathbb{R}^{n_2} \times \mathbb{R}^{l_1}}{\text{minimize}} \quad \langle e_1, p \rangle \quad \text{subject to: } y \geq 0, E_2 y = e_2, \\ & \quad \quad \quad -Ay + E_1^T p \geq 0, \end{aligned} \quad (1)$$

and the dual LCP

$$\begin{aligned} & \underset{(x, q) \in \mathbb{R}^{n_1} \times \mathbb{R}^{l_2}}{\text{maximize}} \quad -\langle e_2, q \rangle \quad \text{subject to: } x \geq 0, E_1 x = e_1, \\ & \quad \quad \quad A^T x + E_2^T q \geq 0. \end{aligned} \quad (2)$$

The vectors $p = (p_0, p_1, \dots, p_{l_2-1}) \in \mathbb{R}^{l_2}$ and $q = (q_0, q_1, \dots, q_{l_1-1}) \in \mathbb{R}^{l_1}$ are dual variables. A is the *payoff matrix* and each E_k is a matrix whose entries are $-1, 0$ or 1 , with exactly 1 entry per row which

equals -1 except for the first whose first entry is 1 and all the others are 0. Each of the vectors e_k is of the form $(1, 0, \dots, 0)$.

The LCPs above have the equivalent saddle-point formulation

$$\underset{y \in Q_2}{\text{minimize}} \quad \underset{x \in Q_1}{\text{maximize}} \quad \langle x, Ay \rangle, \quad (3)$$

where the compact convex polytope

$$Q_k := \{z \in \mathbb{R}^{n_k} \mid z \geq 0, E_k z = e_k\} \subseteq [0, 1]^{n_k} \quad (4)$$

is identified with the strategy profile of player k in the sequence-form representation. At a feasible point (y, p, x, q) for the LCPs, the *duality gap* $\tilde{G}(y, p, x, q)$ is given by¹

$$\begin{aligned} 0 \leq \tilde{G}(y, p, x, q) &:= \langle e_1, p \rangle - (-\langle e_2, q \rangle) = \langle e_1, p \rangle + \langle e_2, q \rangle \\ &= G(x, y) := \max\{\langle u, Ay \rangle - \langle x, Av \rangle \mid (u, v) \in Q_1 \times Q_2\}. \end{aligned} \quad (5)$$

In (5), the quantity $G(x, y)$ is nothing but the primal-dual gap for the equivalent saddle-point problem (3). It was shown (see Theorem 3.14 of [8]) that a pair $(x, y) \in Q_1 \times Q_2$ of realization plans is a solution to the LCPs (1) and (2) (i.e. is a Nash-equilibrium for the game) if and only if there exist vectors p and q such that

$$\begin{aligned} -Ay + E_1^T p &\geq 0, \quad A^T x + E_2^T q \geq 0, \quad \langle x, -Ay + E_1^T p \rangle = 0, \\ \langle y, A^T x + E_2^T q \rangle &= 0. \end{aligned} \quad (6)$$

Moreover, at equilibria *strong duality* holds and the value of the game equals $p_0 = -q_0$, i.e. the duality gap $\tilde{G}(y, p, x, q)$ defined in (5) vanishes at equilibria.

Definition 1 (Nash ϵ -equilibria). Given $\epsilon > 0$, a Nash ϵ -equilibrium is a pair (x^*, y^*) of realization plans for which there exists dual vectors p^* and q^* for problems (1) and (2) such that the duality gap at (y^*, p^*, x^*, q^*) doesn't exceed ϵ . That is,

$$0 \leq \tilde{G}(y^*, p^*, x^*, q^*) \leq \epsilon. \quad (7)$$

1.2 A remark concerning matrix games on simplexes

It should be noted that any matrix $A \in \mathbb{R}^{n_1 \times n_2}$ specifies a matrix game with payoff matrix A , for which player k 's strategy profile is a simplex

$$\Delta_{n_k} := \left\{ z \in \mathbb{R}^{n_k} \mid z \geq 0, \sum_j z_j = 1 \right\}. \quad (8)$$

This simplex can be written as a compact convex polytope in the form (4) by taking $E_k := (1, 1, \dots, 1) \in \mathbb{R}^{1 \times n_k}$ and $e_k = 1 \in \mathbb{R}^1$. Thus every matrix game on simplexes can be seen as a sequential game, and so the results presented in this manuscript can be trivially applied such games in particular. For this special sub-class of sequential games, the duality gap function $G(x, y)$ writes

$$\begin{aligned} G(x, y) &= \max\{\langle u, Ay \rangle - \langle x, Av \rangle \mid (u, v) \in \Delta_{n_1} \times \Delta_{n_2}\} \\ &= \max_{0 \leq i < n_1} (Ay)_i - \min_{0 \leq j < n_2} (A^T x)_j. \end{aligned} \quad (9)$$

¹The first inequality being due to *weak duality*.

1.3 Quick sketch of our contribution

We now give a brief overview of our contributions, which will be made more elaborate in section 3. Developing on an alternative notion of approximate equilibria (see Definition 3) homologous to that presented in Definition 1, we devise a simple numerically stable primal-dual algorithm that (Algorithm 1) for computing approximate Nash-equilibria in sequential two-person zero-sum games with incomplete information and perfect recall. On, each iteration, the only operations performed by our algorithm are of the form $A^T x$, Ay , $E_1^T p$, $E_2^T q$, and $(x)_+ := (\max(0, x_j))_j$. We also prove (Theorem 1) that –in an ergodic / Cesàro sense– the number of iterations required by the algorithm to produce an approximation equilibrium to a precision ϵ is $\mathcal{O}(1/\epsilon)$, with explicit values for the constants involved.

1.4 Notation and terminology

General. Let m and n be positive integers. The components of a vector $z \in \mathbb{R}^n$ will be denoted z_0, z_1, \dots, z_{n-1} (indexing begins from 0, not 1). $\mathbb{R}_+^n := \{z \in \mathbb{R}^n \mid z \geq 0\}$ is the nonnegative *orthant*. $\|z\|$ denotes the 2-norm of z defined by $\|z\| := \sqrt{\langle z, z \rangle}$. Given a matrix $A \in \mathbb{R}^{m \times n}$, its *spectral norm*, denoted $\|A\|$, is defined to be the largest *singular value* of A , i.e the largest *eigenvalue* of $A^T A$ (or equivalently, of AA^T).

Convex analysis. Given a subset $C \subseteq \mathbb{R}^n$, i_C denotes the *indicator function* of C defined by $i_C(x) = 0$ if $x \in C$ and $+\infty$ otherwise. At times, we will write $i_{x \in C}$ for $i_C(x)$ (to ease notation, etc.). For example, we will write $i_{z \geq 0}$ for $i_{\mathbb{R}_+^n}(z)$, etc. The *orthogonal projector* onto C , is the “closest-point” map $\text{proj}_C : \mathbb{R}^n \rightarrow C, x \mapsto \arg\min_{z \in C} \frac{1}{2} \|z - x\|^2$. Let $f : \mathbb{R}^n \rightarrow (-\infty, +\infty]$ be a convex function. The *effective domain* of f , denoted $\text{dom}(f)$, is defined as $\text{dom}(f) := \{x \in \mathbb{R}^n \mid f(x) < +\infty\}$. If $\text{dom}(f) \neq \emptyset$ then we say f is *proper*. The *subdifferential* of f at a point $x \in \mathbb{R}^n$ is defined by $\partial f(x) := \{v \in \mathbb{R}^n \mid f(z) \geq f(x) + \langle v, z - x \rangle, \forall z \in \mathbb{R}^n\}$. If f is convex, its *proximal operator* is the function $\text{prox}_f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ defined by $\text{prox}_f(x) := \arg\min_{z \in \mathbb{R}^n} \frac{1}{2} \|z - x\|^2 + f(z)$.

We recommend [9, 10] for a more detailed exposition on convex analysis and its use in modern optimization theory and practice.

2 Prior work

Here, we present a selection of algorithms that is representative of the efforts that have been made in the literature to compute Nash ϵ -equilibria for two-person zero-sum games with incomplete information like Texas Hold’em Poker, etc. It should be noted that the class of games considered here (sequential games with incomplete information), the LCPs (1) and (2) are exceedingly larger than what state-of-the-art LCP and interior-point solvers can handle (see [11, 12]).

2.1 Regret minimization

CFR (CounterFactual Regret minimization) [1], Monte Carlo CFR [2], and CFR+ [3], by their large popularity, have become the definitive state-of-the-art, and are particularly useful in many-player games, since convex-analytical methods cannot help much in such games. Also, they can be shown to converge to a Nash-equilibrium provided each player uses a CFR scheme to play the game [1], but have a much weaker convergence theory. For example, [2] showed that such schemes have a prohibitive running time of $\mathcal{O}(1/\epsilon^2)$ to produce a Nash ϵ -equilibrium.

2.2 First-order methods

In [11], a nested iterative procedure using the Excessive Gap Technique (EGT) [13] (EGT and precursors are well-known to the signal-processing community [14]) was used to solve the equilibrium problem (3). The authors reported a $\mathcal{O}(1/\epsilon)$ convergence rate (which derives from the general EGT theory) for the outer-most iteration loop. [12] proposed a modified version of the techniques in [11] and proved a $\mathcal{O}((\|A\|/\delta) \log(1/\epsilon))$ convergence rate in terms of the number of calls made to a first-order oracle. Here

$\delta = \delta(A, E_1, E_2, e_1, e_2) > 0$ is a certain *condition number* for the game. The crux of their technique was to observe that (3) can further be written as the minimization of the duality gap function $G(x, y)$ (defined in (5)) for the game², viz

$$\text{minimize} \{G(x, y) \mid (x, y) \in Q_1 \times Q_2\}, \quad (10)$$

and then show there exists a scalar $\delta > 0$ such that for any pair of realization plans $(x, y) \in Q_1 \times Q_2$,

$$\text{“distance between } (x, y) \text{ and set of equilibria”} \leq G(x, y)/\delta. \quad (11)$$

Their algorithm is then derived by iteratively applying Nesterov smoothing [15] with a geometrically decreasing sequence of tolerance levels $\epsilon_{n+1} = \epsilon_n/\gamma$ (with $\gamma > 1$) G . However, there are a number of issues, most notably: (a) The constant $\delta > 0$ can be arbitrarily small, and so the factor $\|A\|/\delta$ in the $\mathcal{O}((\|A\|/\delta) \log(1/\epsilon))$ convergence rate can be arbitrarily large for ill-conditioned games. (b) The reported linear convergence rate is not in terms of basic operations (addition, multiplication, matvec, clipping, etc.), but in terms of the number of calls to a first-order oracle. Most notably, projections onto the polytopes Q_k are computed on each iteration, a very hard sub-problem.

Recently, [16] proposed accelerations to first-order methods for computing Nash-equilibria (including those just discussed), by an appropriate choice of the underlying *Bregman distance* and the *distance generating function* (essential ingredients in EGT-type algorithms). These modifications provably gain a constant factor in the worst-case convergence rate over the original algorithm.

2.3 Primal-dual algorithms

The primal-dual algorithm first developed in [17], was proposed in [18] as a way of solving matrix games on simplexes. Notably, such matrix games on simplexes are considerably simpler than the games considered here. Indeed, the authors in [18] used the fact that computing the orthogonal projection of a point onto a simplex can be done in linear time as in [19]. In contrast, no such efficient algorithm is known nor is likely to exist, for the polytopes Q_k defined in (4). That notwithstanding, such projections can still be done iteratively using for example, the algorithm in proposition 4.2 of [20] or the algorithms developed in [21]. Unfortunately, as with any nested iterative scheme, one would have to solve this sub-problem with finer and finer precision, rendering the overall solver impractical. One can also cite [22], in which the authors endeavored an iterative projection algorithm onto polytopes in outer representation.

Other than the difficult projection sub-problem just discussed, the duality gap might explode even at points arbitrarily close to the set of feasible points, leaving the algorithm with no indication whatsoever, on whether progress is being made.

3 Our contributions

3.1 Generalized Saddle-point Problem (GSP) equivalent for Nash-equilibrium LCPs

In the next theorem, we show that the LCPs (1) and (2) can be conveniently written as a GSP in the sense of [23]. The crux of idea is to remove the linear constraints in the definitions of the strategy polytopes Q_k , by augmenting the payoff matrix to yield an equivalent saddle-point problem. The result is an equivalent game with unbounded strategy profiles (nonnegative orthants) with much simpler geometry. We elaborate the construction in the following theorem.

Theorem 1. Define two proper closed convex functions

$$\left. \begin{aligned} g_1 : \mathbb{R}^{n_2} \times \mathbb{R}^{l_1} &\rightarrow (-\infty, +\infty], & g_1(y, p) &:= i_{y \geq 0} + \langle e_1, p \rangle \\ g_2 : \mathbb{R}^{n_1} \times \mathbb{R}^{l_2} &\rightarrow (-\infty, +\infty], & g_2(x, q) &:= i_{x \geq 0} + \langle e_2, q \rangle \end{aligned} \right\} \quad (12)$$

Also define two bilinear forms $\Psi_1, \Psi_2 : \mathbb{R}^{n_2} \times \mathbb{R}^{l_1} \times \mathbb{R}^{n_1} \times \mathbb{R}^{l_2} \rightarrow \mathbb{R}$ by letting

$$K := \begin{bmatrix} A & -E_1^T \\ E_2 & 0 \end{bmatrix}, \quad \Psi_1(y, p, x, q) := \left\langle \begin{bmatrix} x \\ q \end{bmatrix}, K \begin{bmatrix} y \\ p \end{bmatrix} \right\rangle, \quad (13)$$

²The minimizers of G are precisely the equilibria of the game.

with $\Psi_2 = -\Psi_1$, and define the functions $\hat{\Psi}_1, \hat{\Psi}_2: \mathbb{R}^{n_2} \times \mathbb{R}^{l_1} \times \mathbb{R}^{n_1} \times \mathbb{R}^{l_2} \rightarrow (-\infty, +\infty]$ by

$$\begin{aligned}\hat{\Psi}_1(y, p, x, q) &:= \begin{cases} \Psi_1(y, p, x, q) + g_1(y, p), & \text{if } y \geq 0, \\ +\infty, & \text{otherwise} \end{cases} \\ \hat{\Psi}_2(y, p, x, q) &:= \begin{cases} \Psi_2(y, p, x, q) + g_2(x, q), & \text{if } x \geq 0, \\ +\infty, & \text{otherwise.} \end{cases}\end{aligned}\quad (14)$$

Finally, define the sets $S_1 := \mathbb{R}_+^{n_2} \times \mathbb{R}^{l_1}$ and $S_2 := \mathbb{R}_+^{n_1} \times \mathbb{R}^{l_2}$, and consider the GSP(Ψ_1, Ψ_2, g_1, g_2): Find a quadruplet $(y^*, p^*, x^*, q^*) \in S_1 \times S_2$ s.t. $\forall (y, p, x, q) \in S_1 \times S_2$, we have

$$\begin{aligned}\hat{\Psi}_1(y^*, p^*, x^*, q^*) &\leq \hat{\Psi}_1(y, p^*, x^*, q^*), \quad \text{and} \\ \hat{\Psi}_2(y^*, p^*, x^*, q^*) &\leq \hat{\Psi}_2(y^*, p^*, x, q^*).\end{aligned}\quad (15)$$

Then GSP(Ψ_1, Ψ_2, g_1, g_2) is equivalent to the LCPs (1) and (2), i.e. a quadruplet $(y^*, p^*, x^*, q^*) \in \mathbb{R}^{n_2} \times \mathbb{R}^{l_1} \times \mathbb{R}^{n_1} \times \mathbb{R}^{l_2}$ solves the LCPs (1) and (2) iff it solves GSP(Ψ_1, Ψ_2, g_1, g_2).

Proof. It suffices to show that at any point $(y, p, x, q) \in S_1 \times S_2$, the duality gap between the primal LCP (1) and the dual LCP (2) equals the duality gap of GSP(Ψ_1, Ψ_2, g_1, g_2). Indeed, the unconstrained objective in (1), say $a(x, y)$, can be computed as

$$\begin{aligned}a(y, p) &= \langle e_1, p \rangle + i_{y \geq 0} + i_{-Ay + E_1^T p \geq 0} + i_{E_2 y = e_2} \\ &= g_1(y, p) + \max_{x' \geq 0} \langle x', Ay - E_1^T p \rangle + \max_{q'} \langle q', E_2 y - e_2 \rangle \\ &= g_1(y, p) + \max_{x', q'} \langle x', Ay \rangle - \langle x', E_1^T p \rangle + \langle q', E_2 y \rangle \\ &\quad - (i_{x' \geq 0} + \langle e_2, q \rangle) \\ &= g_1(y, p) - \min_{x', q'} \Psi_2(y, p, x', q') + g_2(x', q') \\ &= g_1(y, p) - \underbrace{\min_{x', q'} \hat{\Psi}_2(y, p, x', q')}_{\phi_2(y, p)} = g_1(y, p) - \phi_2(y, p).\end{aligned}$$

Similarly, the unconstrained objective, say $b(x, q)$, in the dual LCP (2) writes

$$\begin{aligned}b(x, q) &= -\langle q, e_2 \rangle - i_{x \geq 0} - i_{A^T x + E_2^T q \geq 0} - i_{E_1 x = e_1} \\ &= -g_2(x, q) + \min_{y' \geq 0} \langle y', A^T x + E_2^T q \rangle + \min_{p'} \langle p', e_1 - E_1 x \rangle \\ &= -g_2(x, q) + \min_{y', p'} \Psi_1(y', p', x, q) + g_1(y', p') \\ &= -g_2(x, q) + \underbrace{\min_{y', p'} \hat{\Psi}_1(y', p', x, q)}_{\phi_1(x, q)} = -g_2(x, q) + \phi_1(x, q).\end{aligned}$$

Thus, noting that $-\infty < \phi_1(x, q), \phi_2(y, p) < +\infty$ (so that all the operations below are valid), one computes the duality gap between the primal LCP (1) and dual the LCP (2) at (y, p, x, q) as

$$\begin{aligned}a(y, p) - b(x, q) &= g_1(y, p) - \phi_2(y, p) + g_2(x, q) - \phi_1(x, q) \\ &= \Psi_1(y, p, x, q) + g_1(y, p) - \phi_2(y, p) + \Psi_2(y, p, x, q) + g_2(x, q) \\ &\quad - \phi_1(x, q)\end{aligned}$$

$$= \hat{\Psi}_1(y, p, x, q) - \phi_1(x, q) + \hat{\Psi}_2(y, p, x, q) - \phi_2(y, p)$$

$$= \text{duality gap of GSP}(\Psi_1, \Psi_2, g_1, g_2) \text{ at } (y, p, x, q),$$

where the second equality follows from $\Psi_1 + \Psi_2 = 0$. \square

By Theorem 1, solving for a Nash-equilibrium for the game is equivalent to solving the GSP (15), which as it turns out, is simpler conceptually (e.g. we no longer need to compute the complicated orthogonal projections proj_{Q_k}). The rest of the paper will be devoted to developing an algorithm for solving the latter.

3.2 The proposed algorithm

We now derive the algorithm (Algorithm 1) for computing Nash $(\epsilon, 0)$ -equilibria and establish its theoretical properties. The algorithm, which emerges as a synthesis of Theorem 1 above and ideas from [23], is numerically stable and performs only basic iterations (i.e. matvec multiplications, clipping, etc., and no calls to external first-order oracles, no matrix inversions, etc.).

Definition 2. Given $\epsilon > 0$ and a function $f: \mathbb{R}^n \rightarrow (-\infty, +\infty]$, the ϵ -enlarged subdifferential (or ϵ -subdifferential, for short) of f is the set-valued function defined by

$$\partial_\epsilon f(x) := \{v \in \mathbb{R}^n \mid f(z) \geq f(x) + \langle v, z - x \rangle - \epsilon, \forall z \in \mathbb{R}^n\}. \quad (16)$$

The idea behind ϵ -subdifferentials is the following. Say we wish to minimize a convex function f . Replace the usual necessary and sufficient condition “ $0 \in \partial f(x)$ ” for the optimality of x with the weaker condition “ $\partial_\epsilon f(x)$ contains a sufficiently small vector v ”. This approximation concept for subdifferentials yields yet another notion of approximate Nash-equilibrium. the following concept of approximate Nash-equilibria (refer to [23]), namely

Definition 3 (Nash (ϵ_1, ϵ_2) -equilibria). Given tolerance levels $\epsilon_1, \epsilon_2 > 0$, a Nash (ϵ_1, ϵ_2) -equilibrium for the GSP (15) is any quadruplet (x^*, y^*, x^*, q^*) for which there exists a perturbation vector v^* such that $\|v^*\| \leq \epsilon_1$ and $v^* \in \partial_{\epsilon_2}[\hat{\Psi}_1(\cdot, \cdot, x^*, q^*) + \hat{\Psi}_2(y^*, p^*, \cdot, \cdot)](y^*, p^*, x^*, q^*)$. Such a vector v^* is called a Nash (ϵ_1, ϵ_2) -residual at the point (x^*, y^*, x^*, q^*) .

The above definition is a generalization of the notion of Nash-equilibria since: (a) exact Nash-equilibria correspond to Nash $(0, 0)$ -equilibria, and (b) Nash ϵ -equilibria (in the sense of Definition 1) correspond to Nash $(0, \epsilon)$ -equilibria.

Algorithm 1 Primal-dual algorithm for computing Nash $(\epsilon, 0)$ -equilibria in two-person zero-sum sequential games

Require: $\epsilon > 0$; $(y^{(0)}, p^{(0)}, x^{(0)}, q^{(0)}) \in \mathbb{R}^{n_2} \times \mathbb{R}^{l_1} \times \mathbb{R}^{n_1} \times \mathbb{R}^{l_2}$.
Ensure: A Nash $(\epsilon, 0)$ -equilibrium $(y^*, p^*, x^*, q^*) \in S_1 \times S_2$ for the GSP (15).

```

1: Initialize:  $\lambda \leftarrow 1/\|K\|, v^{(0)} \leftarrow 0, k \leftarrow 0$ 
2: while  $k = 0$  or  $\frac{1}{k\lambda}\|v^{(k)}\| \geq \epsilon$  do
3:    $y^{(k+1)} \leftarrow (y^{(k)} - \lambda(A^T x^{(k)} + E_2^T q^{(k)}))_+, p^{(k+1)} \leftarrow$ 
      $p^{(k)} - \lambda(e_1 - E_1 x^{(k)})$ 
4:    $x^{(k+1)} \leftarrow (x^{(k)} + \lambda(Ay^{(k+1)} - E_1^T p^{(k+1)}))_+, \Delta x^{(k+1)} \leftarrow$ 
      $x^{(k+1)} - x^{(k)}$ 
5:    $\Delta q^{(k+1)} \leftarrow \lambda(E_2 y - e_2), q^{(k+1)} \leftarrow q^{(k)} + \Delta q^{(k+1)}$ 
6:    $y^{(k+1)} \leftarrow y^{(k+1)} - \lambda(A^T \Delta x^{(k+1)} + E_2^T \Delta q^{(k+1)}), \Delta y^{(k+1)} \leftarrow$ 
      $y^{(k+1)} - y^{(k)}$ 
7:    $p^{(k+1)} \leftarrow p^{(k+1)} + \lambda E_1 \Delta x^{(k+1)}, \Delta p^{(k+1)} \leftarrow p^{(k+1)} -$ 
      $p^{(k)}$ 
8:    $v^{(k+1)} \leftarrow v^{(k)} + (\Delta y^{(k+1)}, \Delta p^{(k+1)}, \Delta x^{(k+1)}, \Delta q^{(k+1)})$ 
9:    $k \leftarrow k + 1$ 
10: end while
```

Theorem 2 (Ergodic / Cesàrio $\mathcal{O}(1/\epsilon)$ convergence). Let d_0 be the euclidean distance between the starting point $(y^{(0)}, p^{(0)}, x^{(0)}, q^{(0)})$ of Algorithm 1 and the set of equilibria for the GSP (15). Then given any $\epsilon > 0$, there exists an index $k_0 \leq \frac{2d_0\|K\|}{\epsilon}$ such that after k_0 iterations the algorithm produces a quadruplet $(y^{k_0}, p^{k_0}, x^{k_0}, q^{k_0})$ and a vector v^{k_0} such that $\|v^{k_0}\| \leq \epsilon$ and $v^{k_0} \in \partial[\hat{\Psi}_1(\cdot, \cdot, x^{k_0}, q^{k_0}) + \hat{\Psi}_2(y^{k_0}, p^{k_0}, \cdot, \cdot)](y^{k_0}, p^{k_0}, x^{k_0}, q^{k_0})$, where

$$v_a^{(k_0)} := \frac{1}{k\lambda} v^{(k_0)}. \quad (17)$$

Thus Algorithm 1 outputs a Nash $(\epsilon, 0)$ -equilibrium for the GSP (15) in at most $\frac{2d_0\|K\|}{\epsilon}$ iterations.

Proof. It is clear that the quadruplet $(\Psi_1, \Psi_2, g_1, g_2)$ satisfies assumptions B.1, B.2, B.3, B.5, and B.6 of [23] with $L_{xx} = L_{yy} = 0$ and $L_{xy} = L_{yx} = \|K\|$. Now, one easily computes the proximal operator of g_j in closed-form as $\text{prox}_{\lambda g_j}(a, b) \equiv ((a)_+, b - \lambda e_j)$. With all these ingredients in place, Algorithm 1 is then obtained from [23, Algorithm T-BD] applied on the GSP (15) with the choice of parameters: $\sigma = 1 \in (0, 1]$, $\sigma_x = \sigma_y = 0 \in [0, \sigma)$, $\lambda_{xy} := \frac{1}{\sigma L_{xy}} \sqrt{(\sigma^2 - \sigma_x^2)(\sigma^2 - \sigma_y^2)} = \sigma / \|K\| = 1 / \|K\|$, and $\lambda = \lambda_{xy} \in (0, \lambda_{xy}]$. The convergence result then follows immediately from [23, Theorem 4.2]. \square

3.3 Practical considerations

Efficient computation of Ay and $A^T x$. In Algorithms 1, most of the time is spent pre-multiplying vectors by A and A^T . For *flop-type* Poker games like Texas Hold'em and Rhode Island Hold'em, A (and thus A^T too) is very big (up 10^{14} rows and columns!) but is sparse and has a rich block-diagonal structure (each block is itself the Kronecker product of smaller matrices) which can be carefully exploited, as in [11]. Also the sampling strategies presented in the recent work [16] (section 6), for generating unbiased estimates of Ay and $A^T x$ would readily convert Algorithm 1 into an online and much scalable solver.

Computing $\|K\|$. A major ingredient in the proposed algorithm is $\|K\|$, the 2-norm of the huge matrix K . This can be efficiently computed using the power iteration. Also since $\|K\|$ is only used in defining the step-size $\lambda := 1 / \|K\|$, it may be possible to avoid computing $\|K\|$ altogether, and instead use a line-search / backtrack strategy (see [24], e.g) for setting λ .

Game abstraction. For many variants of Poker, there has been extensive research in lossy / lossless abstraction techniques (for example [25] and more recently, [26, 27]), wherein strategically equivalent or not-so-different situations in the game tree are lumped together. This can drastically reduce the size of the state space from a player's perspective, and ultimately, the size of the matrices A , E_1 , and E_2 , without significantly deviating much from the true game.

4 Numerical experiments results

We now present some proof-of-concept for the algorithm proposed. Results are presented and commented in Figure 1.

Remark 1. *We have not benchmarked our algorithm against the algorithms proposed in [15] and Gilpin's et al. [12] because implementing them from scratch for such games would require us to compute the complicated projections prox_{Q_k} . We recall that avoiding these projections was one of the goals of the manuscript.*

4.1 Basic test-bed: Matrix games on simplexes

As in [15, 18], we generate a 1000×1000 random matrix whose entries are uniformly identically distributed in the closed interval $[-1, 1]$. The results of the experiments are shown in Figure 1(a).

4.2 Kuhn Poker, a "toy" sequential game

This game is a simplified form of Poker developed by Harold W. Kuhn in [28]. It already contains all the complexities (sequentiality, imperfection of information, etc.) of a full-blown Poker game like Texas Hold'em, but is simple enough to serve as a proof-of-concept for the ideas developed in this manuscript. The deck includes only three playing cards: a King, Queen, and Jack. One card is dealt to each player, then the first player must bet or pass, then the second player may bet or pass. If any player chooses to bet the opposing player must bet as well ("call") in order to stay in the round. After both players pass or bet, the player with the highest card wins the pot. The pair of vectors $(x^*, y^*) \in \mathbb{R}^{13+13}$ given by

$$x^* = [1, .759, .759, 0, .241, 1, .425, .575, 0, .275, 0, .275, .725]^T, \\ y^* = [1, 1, 0, .667, .333, .667, .333, 1, 0, 0, 1, 0, 1]^T$$

is a Nash $(10^{-4}, 0)$ -equilibrium computed in 1500 iterations of Algorithm 1. The convergence curves are shown in Fig 1. One easy checks that this equilibrium is feasible. Indeed, one computes

$$E_1 x^* - e_1 = [4.76 \times 10^{-5}, -1.91 \times 10^{-5}, 5.67 \times 10^{-5}, 8.23 \times 10^{-6}, 2.90 \times 10^{-5}, -8.62 \times 10^{-7}, -1.96 \times 10^{-5}]^T \text{ and } E_2 y^* - e_2 = [-7.04 \times 10^{-7}, 2.27 \times 10^{-6}, -3.29 \times 10^{-6}, -1.50 \times 10^{-6}, 2.92 \times 10^{-6}, -4.97 \times 10^{-7}, -5.85 \times 10^{-7}]^T.$$

Finally, one checks that $x^{*T} A y^* = -0.05555$, which agrees to 5 d.p with the value of $-1/18$ computed analytically by H. W. Kuhn in his 1950 paper [28]. The evolution of the dual gap and the expected value of the game across iterations are shown in Figure 1. The results of the experiments are shown in Figure 1(b).

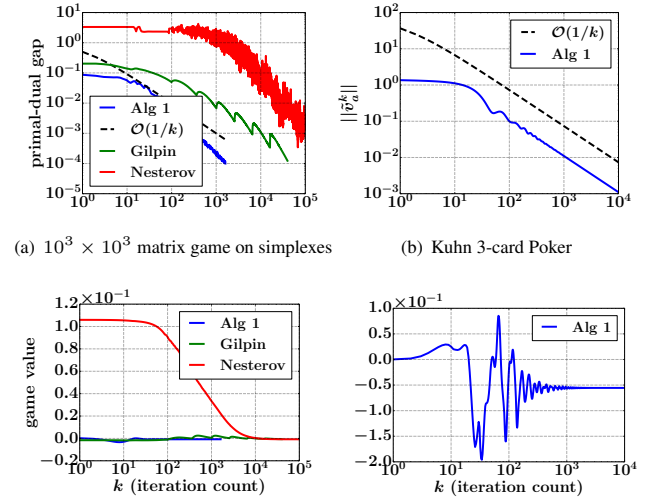


Fig. 1: Convergence curves of Algorithm 1. We stress that the algorithms of Nesterov [15] and Gilpin [12] are included in the plots only indicatively, since this is not meant to be a benchmark as already explained (Remark 1). In (a), the duality gaps are computed according to formula (9). One can see the linear (i.e exponentially fast) behavior of the algorithm in [12], inbetween consecutive breakpoints on the ϵ grid (though the rate of linear convergence seems to be quite close to 1 here). As expected, the first-order smoothing algorithm labelled "Nesterov" [15] jitters around as the iterations go on because even the smoothed problem becomes heavily ill-conditioned near solutions. (b): Kuhn Poker. In the top-right plot, we show the modified duality gap defined in (17). In both cases, we see that the proved convergence rate for our algorithm is empirically observed.

5 Concluding remarks and future work

Making use of the sequence-form representation [5, 7, 8], we have devised a simple numerically stable primal-dual algorithm for computing Nash-equilibria in two-person zero-sum sequential games with incomplete information (like Texas Hold'em, etc.). Our algorithm is simple to implement, with a low constant cost per iteration, and enjoys a rigorous convergence theory with a proven $\mathcal{O}(1/\epsilon)$ convergence in terms of basic operations (matvec products, clipping, etc.), to a Nash $(\epsilon, 0)$ -equilibrium of the game. In future, we plan to run more experiments on real Poker games to measure the practical power of the proposed algorithm compared to other competed schemes like CFR and EGT.

In conclusion, Nash-equilibrium problems are saddle-point convex-concave problems, and as such, a natural tool for tackling them would be proximal primal-dual / operator-splitting algorithms, and we believe such methods will receive more attention in the algorithmic game theory community in future.

6 References

- [1] Martin Zinkevich, Michael Johanson, Michael Bowling, and Carmelo Piccione, “Regret minimization in games with incomplete information,” *Advances in Neural Information Processing Systems*, vol. 20, pp. 1729–1736, 2008.
- [2] Marc Lanctot, Kevin Waugh, Martin Zinkevich, and Michael Bowling, “Monte carlo sampling for regret minimization in extensive games,” in *Advances in Neural Information Processing Systems*, 2009, pp. 1078–1086.
- [3] Michael Bowling, Neil Burch, Michael Johanson, and Oskari Tammelin, “Heads-up limit holdem poker is solved,” *Science*, vol. 347, no. 6218, pp. 145–149, 2015.
- [4] Oskari Tammelin, “Solving large imperfect information games using CFR+,” *CoRR*, vol. abs/1407.5042, 2014.
- [5] Daphne Koller and Nimrod Megiddo, “The complexity of two-person zero-sum games in extensive form,” *Games and economic behavior*, vol. 4, no. 4, pp. 528–552, 1992.
- [6] Daphne Koller, Nimrod Megiddo, and Bernhard Von Stengel, “Fast algorithms for finding randomized strategies in game trees,” in *Proceedings of the twenty-sixth annual ACM symposium on Theory of computing*. ACM, 1994, pp. 750–759.
- [7] Bernhard von Stengel, “Efficient computation of behavior strategies,” *Games and Economic Behavior*, vol. 14, no. 2, pp. 220–246, 1996b.
- [8] Bernhard von Stengel, *Equilibrium Computation for Two-player Games in Strategic and Extensive Form*. N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani (eds.), *Algorithmic Game Theory*, Cambridge Univ. Press, Cambridge UK.
- [9] R. Rockafellar, *Convex analysis (1970)*, Princeton University Press, Princeton, N.J., 1997.
- [10] Patrick L. Combettes and Jean-Christophe Pesquet, “Proximal splitting methods in signal processing,” in *Fixed-point algorithms for inverse problems in science and engineering*, pp. 185–212. Springer New York, 2011.
- [11] Samid Hoda, Andrew Gilpin, Javier Pena, and Tuomas Sandholm, “Smoothing techniques for computing nash equilibria of sequential games,” *Mathematics of Operations Research*, vol. 35, no. 2, pp. 494–512, 2010.
- [12] Andrew Gilpin, Javier Pea, and Tuomas Sandholm, “First-order algorithm with $\mathcal{O}(\ln(1/\epsilon))$ convergence for ϵ -equilibrium in two-person zero-sum games,” *Mathematical Programming*, vol. 133, no. 1-2, pp. 279–298, 2012.
- [13] Yu Nesterov, “Excessive gap technique in nonsmooth convex minimization,” *SIAM Journal on Optimization*, vol. 16, no. 1, pp. 235–249, 2005.
- [14] Jérôme Bobin, Stephen Becker, and Emmanuel Candes, “A fast and accurate first-order method for sparse recovery,” *SIAM J Imaging Sciences*, 2011.
- [15] Yu Nesterov, “Smooth minimization of non-smooth functions,” *Math. Program.*, vol. Ser. A 103, pp. 127152, 2005.
- [16] Christian Kroer, Kevin Waugh, Fatma Kilinc-Karzan, and Tuomas Sandholm, “Faster first-order methods for extensive-form game solving,” in *EC’15*, 2015.
- [17] Antonin Chambolle and Thomas Pock, “A first-order primal-dual algorithm for convex problems with applications to imaging,” *J. Math. Imaging Vis.*, vol. 40, pp. 120, 2011.
- [18] Antonin Chambolle and Thomas Pock, “On the ergodic convergence rates of a first-order primal-dual algorithm,” 2014.
- [19] John Duchi, Shai Shalev-Shwartz, Yoram Singer, and Tushar Chandra, “Efficient projections onto the l_1 -ball for learning in high dimensions,” in *Proceedings of the 25th international conference on Machine learning*. ACM, 2008, pp. 272–279.
- [20] Patrick L. Combettes, Dinh Dũng, and Bng Công Vũ, “Dualization of signal recovery problems,” *Set-Valued and Variational Analysis*, vol. 18, no. 3-4, pp. 373–404, 2010.
- [21] Quoc Tran-Dinh and Volkan Cevher, “Splitting the smoothed primal-dual gap: Optimal alternating direction methods,” *arXiv preprint arXiv:1507.03734*, 2015.
- [22] E.A. Nurminski, “Projection onto polyhedra in outer representation,” *Computational Mathematics and Mathematical Physics*, vol. 48, no. 3, pp. 367–375, 2008.
- [23] Yunlong He and Renato DC Monteiro, “Accelerating block-decomposition first-order methods for solving generalized saddle-point and nash equilibrium problems,” *Optimization-online preprint*, 2013.
- [24] Brendan ODonoghue and Emmanuel Candes, “Adaptive restart for accelerated gradient schemes,” *Foundations of computational mathematics*, vol. 15, no. 3, pp. 715–732, 2013.
- [25] Andrew Gilpin and Tuomas Sandholm, “Lossless abstraction of imperfect information games,” *J. ACM*, vol. 54, 2007.
- [26] Tuomas Sandholm, “Abstraction for solving large incomplete-information games,” *AAAI Conference on Artificial Intelligence (AAAI). Senior Member Track*, 2015.
- [27] Noam Brown, Sam Ganzfried, and Tuomas Sandholm, “Hierarchical abstraction, distributed equilibrium computation, and post-processing, with application to a champion no-limit texas hold’em agent,” in *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, 2015, pp. 7–15.
- [28] H. W. Kuhn, “A simplified two-person poker,” *Annals of Math*, 1950.